UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/733,750	12/11/2003	George Henry Forman	200308713-1	8407

22879 7590 09/27/2006

HEWLETT PACKARD COMPANY
P O BOX 272400, 3404 E. HARMONY ROAD
INTELLECTUAL PROPERTY ADMINISTRATION
FORT COLLINS, CO 80527-2400

EXAMINER	
LE, MICHAEL	

ART UNIT	PAPER NUMBER
2163	

DATE MAILED: 09/27/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/733,750	FORMAN ET AL.	
	Examiner	Art Unit	
	Michael Le	2163	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *11 December 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-23* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-23* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *11 December 2003* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Summary and Status of Claims

1. This Office Action is in response to Application No. 10/733,750 filed December 11, 2003.

2. Claims 1-23 are pending.

3. Claims 6, 13 and 15 are objected to for minor informalities.

4. Claims 2-4 and 22 are rejected under 35 U.S.C. 112, first paragraph.

5. Claims 1-14, 18 and 21-23 are rejected under 35 U.S.C. 112, second paragraph.

6. Claims 1-8 are rejected under 35 U.S.C. 101.

7. Claims 1, 5, 7 and 8 and 23 are rejected under 35 U.S.C. 102(b) as being anticipated by Haimowitz et al. (US Patent 5,819,291).

8. Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291), in view of Mori et al. (US Patent 5,806,058).

9. Claims 9-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291), in view of Commons et al. (US Patent Pub 2002/0069195), further in view of Kuga et al. (US Patent 5,276,616).

10. Claims 15, 18, 19 and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291), in view of Commons et al. (US Patent Pub 2002/0069195).

11. Claims 16 and 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291), in view of Commons et al. (US Patent Pub 2002/0069195), further in view of Kuga et al. (US Patent 5,276,616).

12. Claim 20 is rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291), in view of Commons et al. (US Patent Pub 2002/0069195), further in view of Kuga et al. (US Patent 5,276,616), further in view of Vagnozzi (US Patent 6,070,164).

Priority

13. Claims 1-23 have been examined with a priority date of December 11, 2003.

Drawings

14. The drawings are objected to because in figure 2, reference character "209" has to be changed to --207--. Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as "amended." If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Specification

15. Applicant is reminded of the proper language and format for an abstract of the disclosure.

The abstract should be in narrative form and generally limited to a single paragraph on a separate sheet within the range of 50 to 150 words. It is important that the abstract not exceed 150 words in length since the space provided for the abstract on the computer tape used by the printer is limited. The form and legal phraseology often used in patent claims, such as "means" and "said," should be avoided. The abstract should describe the disclosure sufficiently to assist readers in deciding whether there is a need for consulting the full patent text for details.

The language should be clear and concise and should not repeat information given in the title. It should avoid using phrases which can be implied, such as, "The disclosure concerns," "The disclosure defined by this invention," "The disclosure describes," etc.

16. The abstract of the disclosure is objected to because the language is legal phraseology and in addition, uses implied language. Correction is required. See MPEP § 608.01(b).

17. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

The following title is suggested: System and Method of Retrieving Full Data Through Iterative Cleaning and Matching Techniques.

18. The disclosure is objected to because of the following informalities:

19. On page 2, line 13 --a-- has to be inserted after “closer to”.

20. On page 7, line 5 “recognize” has to be changed to --recognized--.

21. On page 15, line 20 --be-- has to be inserted after “regularly”.

22. On page 17, line 11 --with-- has to be inserted after “accordance”.

23. On page 17, line 15 --be-- has to be inserted after “may”.

24. On page 18, line 20 “of” has to be deleted.

25. Applicant is requested to make any additional corrections that are required upon review of the Specification.

Appropriate correction is required.

Claim Objections

26. **Claims 6, 13 and 15 are objected to** because of the following informalities:

27. In **claim 6**, line 3, --is-- has to be inserted after "stamp".

28. In **claim 13**, line 3, --record-- has to be inserted after "input data" to maintain consistency.

29. In **claim 13**, lines 5 and 9, "cleaned input data" has to be changed to --cleaned input data set-- to maintain consistency.

30. In **claim 13**, line 7, --file—has to be inserted after "associated clean data".

31. In **claim 15**, lines 11 and 14, "cleaned data record" has to be changed to --input cleaned data record-- to maintain consistency with the language of the claim.

Appropriate correction is required.

Claim Rejections - 35 USC § 112

32. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

33. **Claims 2-4 and 22 are rejected under 35 U.S.C. 112, first paragraph**, as failing to comply with the written description requirement. The claim(s) contains subject matter which

was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

34. **Claim 2** recites "a primary clean key record" in lines 1-2. Nowhere in the Specification is "primary clean key record" recited or reasonably described. Furthermore, the limitation cannot be reasonably interpreted as written as explained below in the rejection of claims 2-4 under 35 U.S.C. 112, second paragraph.

35. **Claim 3** also recites the same limitation as claim 2 and is rejected for the same reasons.

36. **Claim 4** recites "a primary complete clean data file" in lines 1-2. Nowhere in the Specification is "a primary complete clean data file" recited or reasonably described. Furthermore, the limitation cannot be reasonably interpreted as written as explained below in the rejection of claims 2-4 under 35 U.S.C. 112, second paragraph.

37. **Claim 22** recites "extending fields of data fields in said table" in lines 3-4. Nowhere in the Specification is "extending fields of data fields in ... [a displayable] ... table" recited or reasonably described. The limitation cannot be reasonably interpreted given the plain meaning of the language as there are multiple interpretations of "extending" in the art. Consequently, claim 22 is not addressed in the prior art rejections below because the Examiner is unable to interpret the claim to apply prior art.

38. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

39. **Claims 1-14, 16-18 and 21-23 are rejected under 35 U.S.C. 112, second paragraph,** as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

40. **Claim 1** recites "...having clean data records and key records wherein *there is* at least one key record..." (emphasis added) in lines 2-3. It is unclear what is meant by "there is". The limitation could be referring back to the table, the clean data records or the key records. The Examiner suggests deleting "there is" and inserting --is-- after "record" in line 3. The Examiner interprets the claim as suggested.

41. **Claim 1** recites the limitation "an associated said clean data record" in line 4. The language of the limitation brings about confusion as to whether the clean data record is another associated clean data record or is the associated clean data record recited in line 3. The Examiner interprets the limitation as "said associated clean data record".

42. **Claims 2-4** are incomprehensible as written. The language of the claims cannot be reasonably interpreted given plain meaning of the words as known in the art nor can it be interpreted in light of the Specification. In fact, a "primary clean key record" is never recited in the Specification thus necessitating the rejection under 35 U.S.C. 112, first paragraph. Specifically, the indefiniteness begins in claim 2. Claim 2 recites that a "clean data record is a primary clean key record of a plurality of said key records in a set". As noted, the Specification nowhere mentions a "primary clean key record". Parent claim 1 recites in line 2 that a persistent table has "clean data records *and* key records". Use of the conjunctive phrase implies that a clean data record is unique from a key record. However, claim 2 recites a limitation that seems to imply that a clean data record is actually one of the key records. Continuing to claim 3, the

primary clean key is limited to being "a pointer to a complete data file associated with *a clean data record*" (emphasis added). It is unclear whether the emphasized portion refers to an additional clean data record, or to the plurality of clean data records recited in claim 1, or to the primary clean data record. Lastly, claim 4 claims further narrows claim 1 and limits the clean data record of claim 1 to be a "primary complete clean data file". A "primary complete clean data file" is never recited in the Specification and given the plain meaning of the language and the language of claim 1, the Examiner is still unable to reasonably interpret the claim. Numerous other problems exist but have been omitted for brevity. Consequently, claims 2-4 are not addressed in the prior art rejections below because the Examiner is unable to interpret the claims to apply prior art.

43. **Claim 9** recites "said input data record" in line 8; "input data" in line 10; "said input data" in line 11; and "said input data record" in line 12. There is lack of antecedent basis for the limitation in the claim. The Examiner suggests amending "incoming data records" in line 7 to --input data records-- and ensuring that each reference of the limitation thereafter is recited as either --said input data record-- or --the input data record--. The claim will be interpreted as suggested.

44. **Claim 9** also recites "said input data record" in line 12. It is unclear whether "said input data record" refers to the input data record before the iterative cleaning or after the iterative cleaning. If it is the latter, then it should be changed to "said cleaned input data record". The limitation is given the latter interpretation in the prior art rejections below.

45. **Claim 9** also recites "a near-matched indexing record" in line 13. It is unclear whether the limitation refers to the near-matched at least one indexing record recited in line 11 or to an

entirely different indexing record. The limitation is given the former interpretation in the prior art rejections below.

46. **Claim 9** also recites "a so-cleaned input data record" in line 14. It is unclear whether the limitation refers to the cleaned input data record recited in claim 10 or an entirely different cleaned input data record. The limitation is given the former interpretation for the prior art rejection below. The Examiner further suggests the limitation be amended to "said cleaned input data record" in both recitations in line 14 and 16.

47. **Claim 9** also recites "an associated one of said clean data files" in line 15. It is unclear whether the limitation refers to the associated one of said clean data files recited in lines 12-13 or an entire newly associated one of said clean data files. The limitation is given the former interpretation in the prior art rejection below.

48. **Claim 12** recites "said so-cleaned input data record" in lines 1-2. The Examiner suggests amending the limitation to "said cleaned input data record" to maintain consistency with the amendments to the limitation in claim 9 as suggested above.

49. **Claim 13** recites "said not recognizing a match therebetween" in line 3. There is a lack of antecedent basis for the limitation in the claim. Claim 13 narrows the "iteratively cleaning" limitation recited in claim 9, line 10. The limitation recited in claim 13 seems redundant as it is apparent through its dependency on claim 9. The Examiner suggests deleting "upon said not recognizing a match therebetween". For the prior art rejection below, the limitation is interpreted as suggested.

50. **Claim 13** also recites "stopping said comparing and retrieving" in lines 6-7. It is unclear whether only the comparing is stopped or both the comparing and the retrieving is stopped. For

the prior art rejection below, the Examiner interprets the limitation to mean stopping only the comparing and subsequently performing the retrieving step.

51. **Claim 13** also recites "said input data" in line 7. It is unclear whether the limitation refers to the input data record before cleaning or the first cleaned input data set recited in line 4. The limitation is given the latter interpretation for the prior art rejection below.

52. Also in **claim 13**, the limitation recited in lines 13-16 is incomprehensible as written. The Examiner is unable to reasonably ascertain the scope of the limitation. For the prior art rejection below, the Examiner interprets the limitation to be "iteratively repeating said re-cleaning and re-comparing until a predetermined phase of cleaning is reached and no said match therebetween is determined, and storing the most recent re-cleaned and re-compared input data record as a new clean data file".

53. **Claim 14** recites "upon said recognizing a match therebetween, a new indexing record is generated for said new clean data file". There limitation is not possible because a "new clean data file" is not created unless there is no match after a predetermined phase of cleaning has been reached, as recited in claim 13. For the prior art rejection below, the Examiner interprets the limitation as "upon reaching the predetermined phase of cleaning with no match, generating a new indexing record for said new clean data file".

54. **Claim 16** recites "an iteratively-generated respective cleaned data record" in lines 2-3. It is unclear whether the limitation refers to the iteratively-generated respective cleaned data record recited in claim 15 or to an entirely different one. For the prior art rejection below, the Examiner interprets the limitation as referring to the clean data record recited in claim 15, lines 8-9.

55. **Claim 18** recites "said generating" in line 3. It is unclear whether the limitation refers to generating a new crude key or iteratively-generated respective cleaned data record. For the prior art rejections below, the Examiner uses the former interpretation.

56. **Claim 21** recites "said tabular array" in line 1. There is lack of antecedent basis for this limitation in the claim. The Examiner interprets the limitation to be "said tabular format".

57. **Claim 22** recites "said tabular array" in line 1. There is lack of antecedent basis for this limitation in the claim. The Examiner interprets the limitation to be "said tabular format".

58. **Claim 23** recites "potentially" in line 5. The limitation "potentially" is a relative word and one of ordinary skill in the art would be unable to reasonably ascertain the meets and bounds of the limitation. For the prior art rejection below, the Examiner interprets the claim without "potentially".

59. **Claim 23** also recites "said array" in line 7. There is lack of antecedent basis for this limitation in the claim. The Examiner interprets the limitation to be "said tabulation of crude keys".

60. **Claims 5-8, 10, 11 and 17** are rejected because they depend on a rejected claim. Dependent claims contain the limitations of the parent claims and are therefore rejected for the same reasons.

61. The prior art rejections below for claims 1-14, 18 and 21-23 are made as best understood in light of the 35 U.S.C. 112, first and second paragraph rejections addressed above.

Claim Rejections - 35 USC § 101

62. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

63. **Claims 1-8 are rejected under 35 U.S.C. 101** because the claimed invention is directed to non-statutory subject matter.

64. The basis of this rejection is set forth in a test of whether the invention is categorized as a process, machine, manufacture or composition of matter and if the invention produces a useful, concrete and tangible result. Mere ideas in the abstract (i.e., abstract idea, law of nature, natural phenomena) are found to be non-statutory subject matter. For a method claim to pass muster, the recited process must produce a useful, concrete and tangible result.

65. In the present case, **claims 1-8** recite an analysis tool, which is interpreted as a software application. A computer program must be tangibly embodied in order for its functional interrelationships with the machine or computer to be realized, thus allowing it to be categorized in a statutory category of invention. Arguably, software code can be categorized in the statutory category of a process, however, to be categorized as a statutory category of a process, the software must be claimed as a method or process and all the steps of the method must be claimed using positive method steps. This is not the case for the analysis tool of claims 1-8, thus it is nonstatutory.

To expedite a complete examination of the instant application, the claims rejected under 35 U.S.C. 101 (nonstatutory) above are further rejected as set forth below in anticipation of applicant amending these claims to place them within the four statutory categories of invention

Claim Rejections - 35 USC § 102

66. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

67. Claims 1, 5, 7 and 8 and 23 are rejected under 35 U.S.C. 102(b) as being anticipated by Haimowitz et al. (US Patent 5,819,291) hereinafter "Haimowitz".

68. In regards to **claim 1**, Haimowitz discloses a heuristics analysis tool comprising:

a. a persistent table, having clean data records and key records wherein at least one key record is associated with each clean data record, said key record having at least one field of data from the associated said clean data record (Col. 2, lines 63-66; col. 4, lines 65-67; col. 5, lines 1-8)[1]; and

b. associated with said key records, heuristic-based routines for automatically generating said key records from each newly received data record for matching to said clean data record (Col. 3, lines 32-43; col. 4, lines 65-67; col. 5, lines 1-8)[2].

69. In regards to **claim 5**, Haimowitz discloses the tool as set forth in claim 1 further comprising: at least one column recording one or more of said heuristic-based routines that were involved in generating each of said key records (Col. 5, lines 1-4, 50-67; col. 6, lines 6-10)[3].

---

[1] The candidate records (key record) are associated with the existing records (clean data records) in the database.
[2] The candidate records (key records) are associated with hash key routines for generating them (heuristic based routines). The candidate records (key records) are generated for existing data files. Once a new data file is entered, a candidate record is created for it.
[3] The hash key field is interpreted as heuristic based routines because they are functions used to generate candidates (key records).

70. In regards to **claims 7 and 8**, Haimowitz discloses special flags associated with said key records, said flags are a quality factor assigned to each said key record (col. 5, lines 31-34).

71. In regards to **claim 23**, Haimowitz discloses a method of doing business comprising:

a. storing a database of clean data files for each of a plurality of entities (Col. 2, lines 62-66);

b. creating a tabulation of crude keys, each having a pointer to an associated one of said clean data files (Col. 3, lines 34-47; col. 4, lines 65-67; col. 5, lines 1-8)[4];

c. receiving periodically a dirty data record related to at least one entity of said plurality of entities (Col. 3, lines 32-33)[5];

d. comparing said dirty data to said tabulation (Col. 3, lines 37-40; col. 6, lines 11-18)[6];

e. assigning said dirty data to one of said clean data files (Col. 6, lines 20-21)[7]; and

f. creating new clean data files from said dirty data when no pointer substantially matching said dirty data is found during said comparing (Col. 6, lines 19-20)[8].

Claim Rejections - 35 USC § 103

72. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

---

[4] The candidate set is interpreted as a tabulation of crude keys. Each of the candidates is related to an existing entry in the database. This is interpreted as having a pointer to a clean data file.
[5] A new record is received is interpreted as receiving a dirty record related to at least one entity of said plurality of entities.
[6] Matching (comparing) between the new record (dirty data) and each of the candidates (tabulation) is performed.
[7] If a match is found the new record (dirty data) is used to update (assigned) the existing record (clean data file) associated with the matched candidate.
[8] If no match is found, then the new data is used to create a new record in the database.

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

73. **Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291) hereinafter "Haimowitz" in view of Mori et al. (US Patent 5,806,058) hereinafter "Mori".**

74. In regards to **claim 6**, Haimowitz does not expressly disclose a time-stamp with each said key record in the table wherein said time-stamp is indicative of most recent use.

75. Mori discloses an index record (key record) having a time access (time-stamp) field that indicates the time the index record was most recently used (Mori: col. 3, lines 45-47).

76. Haimowitz and Mori are analogous art because they are directed to the same field of endeavor of database management with indices.

77. At the time of the invention, it would have been obvious to one of ordinary skill in the art to modify the tool of Haimowitz by adding a time-stamp with each said key record in the table wherein said time-stamp is indicative of most recent use, as taught by Mori.

78. The motivation for doing so would have been because a timestamp indicating the most recent use of the index key would be useful in determining what indices are no longer useful and could potentially be deleted. An index that has not been accessed for a long time is most likely no longer useful and can be deleted to conserve space and increase the efficiency and speed of the database system (Mori: col. 1, lines 40-47, 59-62).

79. Claims 9-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291) hereinafter "Haimowitz" in view of Commons et al.

(US Patent Pub 2002/0069195) hereinafter "Commons", further in view of Kuga et al. (US Patent 5,276,616) hereinafter "Kuga".

80. In regards to **claim 9**, Haimowitz discloses a data association and cleaning method comprising:

a. storing a plurality of clean data files and, associated with each of said clean data files, at least one indexing record, each said indexing record containing at least one field related to a respective associated clean data file such that said at least one indexing record serves as a pointer to the respective associated said clean data file (Haimowitz: Col. 2, lines 62-66; col. 3, lines 35-37; col. 4, lines 65-67; col. 5, lines 1-8)[9];

b. comparing input data records to the indexing records for obtaining a match, and assigning said input data record to the respective associated said clean data file associated with a matched indexing record (Haimowitz: col. 3, lines 31-43; col. 6, lines 11-21)[10];

c. and upon no match, adding said cleaned input data record as a new clean data file with an associated indexing record therefor (Haimowitz: col. 6, lines 18-21)[11].

81. Haimowitz does not expressly disclose if not obtaining a said match, iteratively cleaning said input data record until at least a near-match between said cleaned input data record and said at least one indexing record is obtained and assigning said cleaned input data record to the one of said clean data files associated with a near-matched indexing record and upon a near match,

---

[9] A database of existing records is stored (clean data files). Associated with the existing records are candidates (indexing record). The candidate set is related to the associated existing record and serves as a method of accessing the clean data file (pointer to respective associated clean data file).

[10] An input data record is matched with each of the candidates (indexing record) to obtain a match. If a match is obtained then the input data record is used to update (assigned) the associated existing record in the database (clean data file) associated with the matched candidate (indexing record).

[11] If no match is found the input data is inserted into the database (adding input data record as a new clean data file).

adding a said cleaned input data record as a new indexing record for the associated one of said clean data files.

82. Commons discloses an iterative process for finding a matching database record (Commons: para. 0058, lines 1-3). Commons further discloses using a search key that is matched against records that are stored in the database (Commons: para. 0058, lines 4-6, 10-11). If a match is not found on the first try, the search is repeated (matching is repeated) with progressively less specific information (iteratively cleaned) (Commons: para. 0060, lines 1-3). This process is repeated until a match is found or it is determined that no matching records exist (Commons: Figs. 3A, 3B; para. 0060-63).

83. Kuga discloses an apparatus for generating an index from input data (Kuga: col. 5, lines 9-11). A matching module matches input data to an existing entry in a dictionary (database) to determine a match. When there is a substantial match or an exact match, an index entry is generated and associated with the existing entry (Kuga: col. 13, lines 26-45).

84. Haimowitz, Commons and Kuga are analogous art because they are directed to the same field of endeavor of data storage and retrieval.

85. At the time of the invention, it would have been obvious to one of ordinary skill in the art to modify the method of Haimowitz by adding the steps of iteratively cleaning said input data record until at least a near-match between said cleaned input data record and said at least one indexing record is obtained and assigning said cleaned input data record to the one of said clean data files associated with a near-matched indexing record upon no match, as taught by Commons and adding a said cleaned input data record as a new indexing record for the associated one of said clean data files upon obtaining a match, as taught by Kuga.

86. The motivation for doing so would have been because iteratively cleaning the input data and performing a matching process to retrieve data is fast and accurate process for retrieving data (Commons: para. 0058, lines 1-3). Furthermore, since the process is iterative, manual intervention following a failed matching step is no longer required and thus provides convenience and speed. The motivation for creating an indexing record upon a match would have been because manual index creation is burdensome and inconsistent (Kuga: col. 1, lines 47-68).

87. In regards to **claim 10**, Haimowitz discloses the method as set forth in claim 9 wherein said storing is in a displayable format (Haimowitz: Fig. 5).

88. In regards to **claim 11**, Haimowitz discloses the method as set forth in claim 10 further comprising: at given intervals, performing a data clean-up on a stored table in said displayable format (Haimowitz: Fig. 5; col. 9, lines 30-32; col. 10, lines 4-6, 9-18)[12].

89. In regards to **claim 12**, Haimowitz discloses the method as set forth in claim 9 wherein upon said adding said cleaned input data record as a new clean data file with an associated indexing record therefor, flagging said new clean data file (Haimowitz: col. 9, lines 58-64)[13].

90. **Claim 13** was addressed above in the rejection to claim 9 as being disclosed by Haimowitz and Commons. Haimowitz and Commons disclose said iteratively cleaning (Commons: para. 0058, lines 1-3) further comprising:

---

[12] An administrator can add additional rules and choose indices at given intervals using the graphical user interface shown in figure 5 (displayable format).

[13] The new customer ID generated for the new record is interpreted as flagging the new entry (new clean data file).

a. cleaning said input data record and storing a first cleaned input data set (Commons: para. 0060, lines 1-3);

b. comparing the first cleaned input data set to each said indexing record (Commons: para. 0060, lines 1-3), and

i. upon recognizing a match therebetween, stopping said comparing and retrieving the associated clean data file for association with said cleaned input data record (Commons: Figs. 3A, 3B; para. 0060-63),

ii. upon not recognizing a match therebetween, re-cleaning said first cleaned input data set, discarding said first cleaned input data set, and storing it as a subsequently cleaned input data set; (Commons: Figs. 3A, 3B; para. 0060-63)

c. re-comparing the subsequently cleaned input data set to said indexing record (Commons: Figs. 3A, 3B; para. 0060-63); and

d. iteratively repeating said re-cleaning and re-comparing until a predetermined phase of cleaning is reached and no said match therebetween is determined (Commons: Figs. 3A, 3B; para. 0060-63), and storing the most recent re-cleaned and re-compared input data record as a new clean data file (Haimowitz: col. 6, lines 18-21).

91. In regards to **claim 14**, Haimowitz discloses the method as set forth in claim 13 wherein upon reaching the predetermined phase of cleaning with no match, generating a new indexing record for said new clean data file (Haimowitz: col. 9, lines 58-64).

92. **Claims 15, 18, 19 and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291) hereinafter "Haimowitz" in view of Commons et al. (US Patent Pub 2002/0069195) hereinafter "Commons".**

93. In regards to **claim 15**, Haimowitz discloses a computer memory (Col. 2, lines 61-63) comprising:

a. computer code means for receiving input data records (Haimowitz: Col. 3, lines 32-33);

b. computer code means for comparing said input data records to a tabular format set of crude keys (Haimowitz: Col. 3, lines 37-40);

c. computer code means for returning a clean key associated with one of said crude keys upon a comparing match (Haimowitz: col. 6, lines 18-20)[14];

d. computer code means for creating a new data file from a said input data record such that said new data file is also a first respective one of said crude keys associated therewith (Haimowitz: col. 6, lines 19-20; col. 9, lines 58-64).

94. Haimowitz does not expressly disclose computer code means for iterative cleaning of said input data records upon a no-match return and storing an iteratively-generated respective cleaned data record therefrom, computer code means for re-comparing said iteratively-generated respective cleaned input data record to said set of crude keys, searching for said match return and computer code means for creating a new data file from a last said iteratively-generated respective

[14] When a match is found, the existing record (clean key) associated with the matched candidate (crude key) is accessed (returning).

cleaned input data record such that said new data file is also a first respective one of said crude keys associated therewith.

95. Commons discloses an iterative process for finding a matching database record (Commons: para. 0058, lines 1-3). Commons further discloses using a search key that is matched against records that are stored in the database (Commons: para. 0058, lines 4-6, 10-11). If a match is not found on the first try, the search is repeated (matching is repeated) with progressively less specific information (iteratively cleaned) (Commons: para. 0060, lines 1-3). This process is repeated until a match is found or it is determined that no matching records exist (Commons: Figs. 3A, 3B; para. 0060-63).

96. Haimowitz and Commons are analogous art because they are directed to the same field of endeavor of data storage and retrieval.

97. At the time of the invention, it would have been obvious to one of ordinary skill in the art to modify the computer memory of Haimowitz by adding computer code means for iterative cleaning of said input data records upon a no-match return and storing an iteratively-generated respective cleaned data record therefrom, computer code means for re-comparing said iteratively-generated respective cleaned input data record to said set of crude keys, searching for said match return and computer code means for creating a new data file from a last said iteratively-generated respective cleaned input data record such that said new data file is also a first respective one of said crude keys associated therewith, as taught by Commons.

98. The motivation for doing so would have been because iteratively cleaning the input data and performing a matching process to retrieve data is fast and accurate process for retrieving data (Commons: para. 0058, lines 1-3). Furthermore, since the process is iterative, manual

intervention following a failed matching step is no longer required and thus provides convenience and speed.

99. In regards to **claim 18**, Haimowitz discloses the computer memory as set forth in claim 17 further comprising: computer code means for displaying in said tabular format said crude keys and heuristic routines employed in said generating (Haimowitz: Col. 10, lines 4-18).

100. In regards to **claim 19**, Haimowitz discloses the computer memory as set forth in claim 15 wherein each of said crude keys has an associated pointer to obtain said associated clean key (Col. 4, lines 65-67; col. 5, lines 1-8)[15].

101. In regards to **claim 21**, Haimowitz discloses the computer memory as set forth in claim 15 wherein said tabular format is a displayable table, comprising: computer code means including heuristic routines for editing said table (Haimowitz: col. 9, lines 30-41).

102. **Claims 16 and 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291) hereinafter "Haimowitz" in view of Commons et al. (US Patent Pub 2002/0069195) hereinafter "Commons", further in view of Kuga et al. (US Patent 5,276,616) hereinafter "Kuga".**

103. In regards to **claim 16**, Commons discloses iterative cleaning as addressed above in the rejection to parent claim 15.

104. Haimowitz and Commons do not expressly disclose computer code means for generating a new crude key from the iteratively-generated respective cleaned data record.

[15] Each of the candidates (crude keys) generated are associated with an existing record (clean key) in the database.

105. Kuga discloses an apparatus for generating an index from input data (Kuga: col. 5, lines 9-11). A matching module matches input data to an existing entry in a dictionary (database) to determine a match. When there is a substantial match or an exact match, an index entry is generated and associated with the existing entry (Kuga: col. 13, lines 26-45).

106. Haimowitz, Commons and Kuga are analogous art because they are directed to the same field of endeavor of data storage and retrieval.

107. At the time of the invention, it would have been obvious to one of ordinary skill in the art to modify the combined computer memory of Haimowitz and Commons by adding computer code means for generating a new crude key from the iteratively-generated respective cleaned data record, as taught by Kuga.

108. The motivation for doing so would have been because automatically creating an index record upon a match the burden of manual index creation and maintains consistency (Kuga: col. 1, lines 47-68).

109. In regards to **claim 17**, Haimowitz discloses the computer memory as set forth in claim 16 wherein said computer code means for generating a new crude key has heuristic routines (Haimowitz: Col. 5, lines 1-4, 50-67; col. 6, lines 6-10)[16].

110. **Claim 20 is rejected under 35 U.S.C. 103(a) as being unpatentable over Haimowitz et al. (US Patent 5,819,291) hereinafter "Haimowitz" in view of Commons et al. (US Patent**

[16] The hash key field is interpreted as heuristic based routines because they are functions used to generate candidates (key records).

Pub 2002/0069195) hereinafter "Commons", further in view of Kuga et al. (US Patent 5,276,616) hereinafter "Kuga", further in view of Vagnozzi (US Patent 6,070,164).

111. In regards to **claim 20**, Haimowitz, Commons and Kuga do not expressly disclose each of said crude keys points to a cleanest one of a plurality of crude keys associated with a clean data file.

112. Vagnozzi discloses data values stored in a database that are associated with fine keys (cleanest key), which are in turn associated with one or more coarse keys (crude keys) (Vagnozzi: col. 3, lines 23-45).

113. Haimowitz, Commons, Kuga and Vangozzi are analogous art because they are directed to the same field of endeavor of database systems using indices.

114. At the time of the invention, it would have been obvious to one of ordinary skill in the art to modify the combined computer memory of Haimowitz, Commons and Kuga by making each of said crude keys points to a cleanest one of a plurality of crude keys associated with a clean data file, as taught by Vagnozzi.

115. The motivation for doing so would have been because it allows for fast query responses by minimizing the number of key indices required (Vagnozzi: col. 2, lines 65-67; col. 3, lines 1-17, 21-23).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael Le whose telephone number is 571-272-7970. The examiner can normally be reached on Mon-Thurs : 9:30am-6pm, Fri: 8am-4:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Don Wong can be reached on 571-272-1834. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

DON WONG
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100

Michael Le
Art Unit 2163
September 18, 2006